M 100 35

THE OPEN UNIVERSITY

Mathematics Foundation Course Unit 35

Topology

The Open University

Mathematics Foundation Course Unit 35

TOPOLOGY

Prepared by the Mathematics Foundation Course Team

Correspondence Text 35

The Open University Press

The Open University Press
Walton Hall, Bletchley, Buckinghamshire

First published 1971. Reprinted 1972, 1973

Printed in Great Britain by
EYRE AND SPOTTISWOODE LIMITED
AT GROSVENOR PRESS PORTSMOUTH

SBN 335 01034 2

Open University courses provide a method of study for independent learners through an integrated teaching system including textual material, radio and television programmes and short residential courses. This text is one of a series that makes up the correspondence element of the Mathematics Foundation Course.

For general availability of supporting material referred to in this text, please write to the Director of Marketing, The Open University, Walton Hall, Bletchley, Buckinghamshire.

Further information on Open University courses may be obtained from the Admissions Office, The Open University, P.O. Box 48, Bletchley, Buckinghamshire.

1.3

Contents

Objectives

The principal objective of this unit is to generalize the definition of a continuous function given in *Unit 7, Sequences and Limits I*, and to reformulate it in terms of wider applicability. In doing this, we introduce the subject of topology.

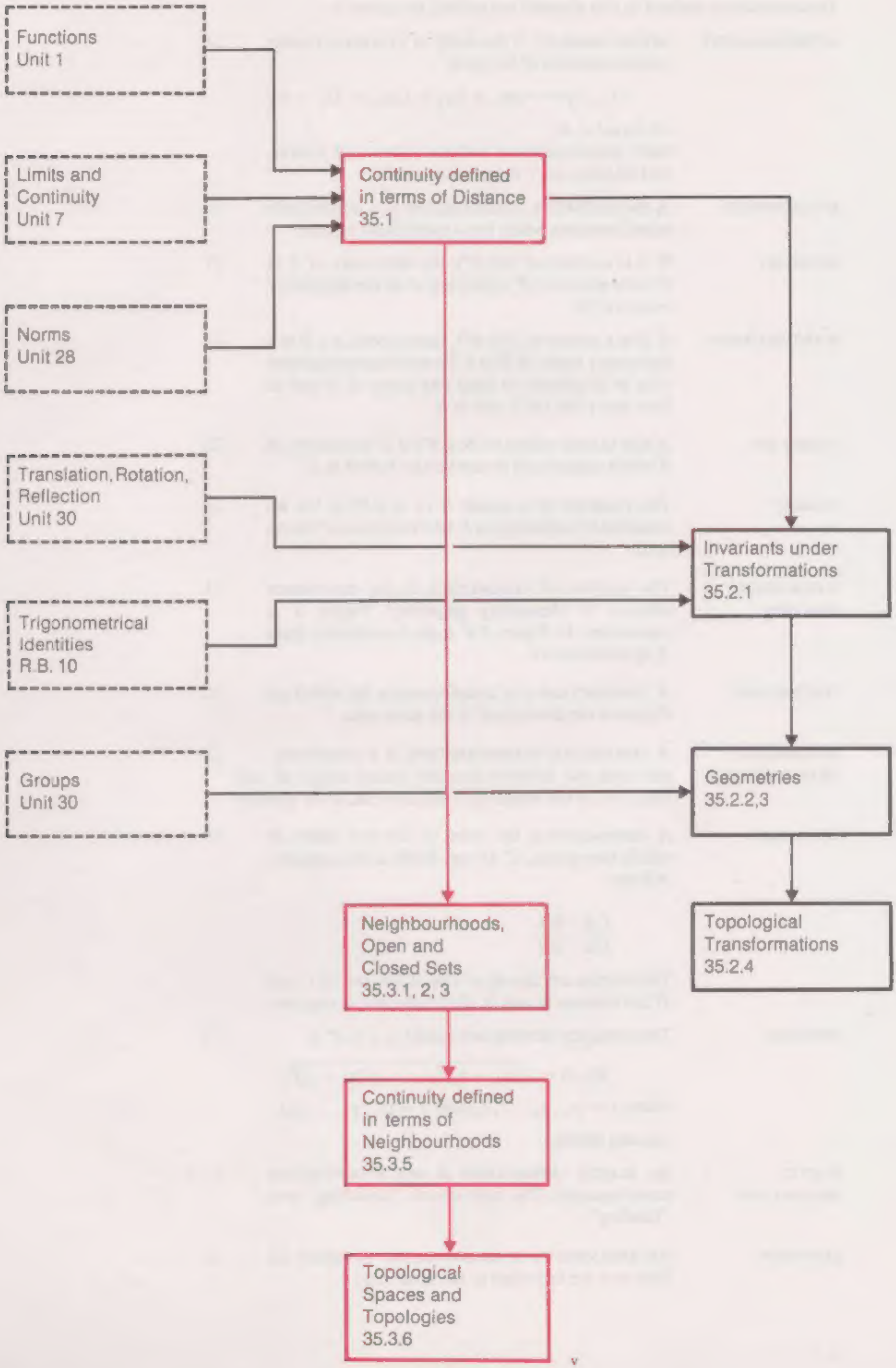
After working through this unit you should be able to:

(i) explain the meanings of the terms: isometry, similarity, topological transformation, neighbourhood, boundary point, boundary, open set, closed set, interior, closure, topological space, a topology;
(ii) describe the neighbourhoods of given points of a set $S \subseteq R^n$;
(iii) find the boundaries of given subsets of $S \subseteq R^n$;
(iv) decide whether given subsets of $S \subseteq R^n$ are open or closed;
(v) define *continuity* in terms of distance, neighbourhoods, open sets.

Note

Before working through this correspondence text, make sure you have read the general introduction to the mathematics course in the Study Guide, as this explains the philosophy underlying the whole course. You should also be familiar with the section which explains how a text is constructed and the meanings attached to the stars and other symbols in the margin, as this will help you to find your way through the text.

Structural Diagram



Glossary

Terms which are defined in this glossary are printed in CAPITALS.

		Page
AFFINE GEOMETRY	AFFINE GEOMETRY is the study of INVARIANTS under transformations of the form $$(x_1, x_2) \longmapsto (ax_1 + bx_2 + l, cx_1 + dx_2 + m),$$ where $ad \neq bc$. Such transformations include SHEAR and STRAIN, and all SIMILARITY TRANSFORMATIONS.	14
BI-CONTINUOUS	A BI-CONTINUOUS transformation is a CONTINUOUS transformation which has a continuous inverse.	15
BOUNDARY	If X is a subset of $S(\subseteq R^n)$, the BOUNDARY of X in S is the subset of R^n consisting of all the BOUNDARY POINTS of X.	21
BOUNDARY POINT	If X is a subset of $S(\subseteq R^n)$, then a point $p \in S$ is a BOUNDARY POINT of X in S if every NEIGHBOURHOOD of p in S includes at least one point of X and at least one point (in S) not in X.	21
CLOSED SET	X is a CLOSED subset of $S(\subseteq R^n)$ if X is a subset of S which includes *all* its BOUNDARY POINTS in S.	22
CLOSURE	The CLOSURE of a subset X of $S(\subseteq R)$ is the set obtained by adjoining to X all its BOUNDARY POINTS in S.	24
CONGRUENCE RELATION	The relation of CONGRUENCE is the equivalence relation of elementary geometry: Figure A is CONGRUENT to Figure B if B can be obtained from A by an ISOMETRY.	11
CONTRACTION	A CONTRACTION is a transformation by which all distances are diminished in the same ratio.	12
CONTINUOUS TRANSFORMATION	A CONTINUOUS TRANSFORMATION is a transformation with the property that the reverse image of each OPEN SET of the image set is an open set in the domain.	28
CROSS-RATIO	A CROSS-RATIO is the ratio of the two ratios in which two points, C, D say, divide a line-segment, AB say: $$\frac{CA}{CB} : \frac{DA}{DB}.$$ The lengths are *signed*, so that if *just one* of C and D lies between A and B, the cross-ratio is negative.	15
DISTANCE	The DISTANCE between two points x, y in R^n is $$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2},$$ where $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$. See also METRIC.	3
ELASTIC DEFORMATION	An ELASTIC DEFORMATION is any BI-CONTINUOUS transformation. The term covers "stretching" and "bending".	1, 15
EXPANSION	An EXPANSION is a transformation by which all distances are increased in the same ratio.	12

		Page
INTERIOR	The INTERIOR of a subset X of $S(\subseteq R^n)$ in S is the set of elements of X which are not BOUNDARY POINTS of X.	24
INVARIANT PROPERTY	An INVARIANT PROPERTY of a particular class of transformations is a property which is preserved by the transformations.	8
ISOMETRY	An ISOMETRY (RIGID TRANSFORMATION) is a one-one transformation which preserves DISTANCE; for example, REFLECTIONS, ROTATIONS and TRANSLATIONS.	11
METRIC	A METRIC is a function of the form $d: R^n \times R^n \longmapsto R$ which defines DISTANCE on R^n with the properties: (i) $d(x, y) \geqslant 0$ (ii) $d(x, y) = 0 \Leftrightarrow x = y$ (iii) $d(x, y) = d(y, x)$ (iv) $d(x, y) + d(y, z) \geqslant d(x, z)$ $(x, y, z \in R^n)$.	19
NEIGHBOURHOOD	For $x \in S(\subseteq R^n)$ and $r > 0$, a NEIGHBOURHOOD of x in S is the set of all points y in S for which $d(x, y) < r$.	19
OPEN SET	X is an OPEN subset of $S(\subseteq R^n)$ if X is a subset of S which contains *none* of its BOUNDARY POINTS in S.	22
REFLECTION	A REFLECTION is an ISOMETRY for which the line-segments joining points to their images all have a common perpendicular bisector.	10
RIGID TRANSFORMATION	See ISOMETRY.	
ROTATION	A ROTATION is an ISOMETRY which leaves just one point of the plane INVARIANT.	9
SHEAR	A SHEAR is a transformation under which a family of parallel lines slip relative to each other whilst preserving the "straightness" of lines crossing them.	14
SIMILARITY RELATION	The relation of SIMILARITY is the equivalence relation: Figure A is SIMILAR to Figure B if B can be obtained from A by a transformation under which straight lines and the angles between them are INVARIANT.	12
SIMILARITY TRANSFORMATION	A SIMILARITY TRANSFORMATION is a one-one transformation which multiplies DISTANCES by a fixed factor $m > 0$.	12
STRAIN	A STRAIN is a transformation by which all distances parallel to some fixed line are changed in the same proportion.	12
TOPOLOGICAL SPACE	A TOPOLOGICAL SPACE consists of a set S together with a collection of its subsets which includes all their unions and all intersections of a finite number of subsets, in addition to the empty set and the set S.	33
TOPOLOGICAL TRANSFORMATION	A TOPOLOGICAL TRANSFORMATION is a one-one BI-CONTINUOUS TRANSFORMATION.	15

		Page
TOPOLOGY	(i) TOPOLOGY is the study of INVARIANTS under TOPOLOGICAL TRANSFORMATIONS. (ii) A TOPOLOGY is a collection of subsets used to define a TOPOLOGICAL SPACE.	1, 15
TORUS	A TORUS is the surface obtainable from a cylinder by bending it around in a circle and joining the two ends: a ring doughnut.	16
TRANSLATION	A TRANSLATION is an ISOMETRY in which all points are moved through the same fixed distance parallel to a given straight line.	8

Notation

The symbols are presented in the order in which they appear in the text.

$d(x, y)$	The distance between x and y, $(x, y \in R^n)$.	3
$\subseteq$	"is a subset of".	5
S	A subset of R^n.	19
$N(x, r, S)$	The neighbourhood of x of radius r in S.	20
$\varnothing$	The empty set.	23
R_0^-	The set of non-positive real numbers.	29
$\mathscr{C}$	A non-empty collection of subsets (the open sets) of a topological space.	33

Bibliography

B. H. Arnold, *Intuitive Concepts in Elementary Topology* (Prentice-Hall, 1962).

This book follows an intuitive approach similar to ours. Chapter 1 discusses the question: "What is Topology?" Chapter 3 deals with topological equivalence in three-dimensional space; Chapter 7 discusses distance and transformations; Chapter 8 deals with topological spaces.

The book contains quite a lot of material not discussed in this text, and should provide interesting and stimulating reading for those who wish to carry the subject further.

35.0 INTRODUCTION

35.0

Introduction

Topology is a fundamental part of present-day mathematics. As a subject to be studied in its own right, it is fairly recent as mathematical subjects go, but it has developed very rapidly in the last few decades. There has been an increasing demand for very precise definitions of what we mean by such terms as "nearness", "continuity", and various properties of mathematical spaces

The word *topology* itself means strictly *the study of position.* It is derived from two Greek words, *τοπος* meaning *place* and *λογια* meaning *study* At one time it was known by the Latin name *analysis situs.* Perhaps nowadays it might be described more exactly as the study of transformations under which certain properties are invariant.

Leonhard Euler

Part of topology is directed towards a study of geometrical objects and how they behave under what have come to be called *elastic deformations*: stretching, bending, but not tearing. But much of topology is a very general and abstract enquiry into the nature of space, and is far removed from geometric intuition. The roots of the specifically geometric aspects of the subject go back to roughly the seventeenth century. Euler, for example, refers to Leibniz as the originator of *analysis situs*, but the first really major attempt at a systematic development of these geometrical aspects was made by Poincaré in 1895.

Gottfried Wilhelm Leibniz

The study of surfaces under elastic deformations has been popularized in various elementary books, and in some films and broadcast material, because certain aspects of it lend themselves to some interesting and sometimes quite spectacular conjuring tricks. However, you should not be misled into thinking that the major part of topology is concerned with whether or not you can take off your vest without removing your jacket, or whether you can turn an inner tube inside out through the hole for its valve. These are tricks of passing interest. We discuss one or two interesting surfaces in the television programme associated with this unit. In this correspondence text we shall introduce some of the formal aspects of topology which enable us to be more precise about some of the words which we tend to use rather glibly — in particular, the word *continuous*

The Mathematics Foundation Course is intended to be a general introduction to what mathematics is and what mathematics does, and it would be incomplete without some reference to this important branch of mathematics, topology. Although this unit comes near the end of the course, you should study the text with care, since it will not only provide a part of the foundation upon which future mathematics courses will be built, but it will also help you to a deeper understanding of some of the topics of earlier units in the course.

A central theme of the text is the concept of a *continuous transformation.* Appeal is made to intuition, but we do give precise formal definitions of a number of fundamental concepts.

In our discussions, we shall give many definitions and results which are applicable to sets in general. However, because our appeals to intuition will be largely of a geometric nature, we shall usually speak in terms of subsets of R, R^2, R^3, and, generally, R^n, and hence we shall speak of *points* $x, y, \ldots \in R^n$, rather than *elements* of some general set.

Henri Poincaré

35.1 CONTINUOUS FUNCTIONS

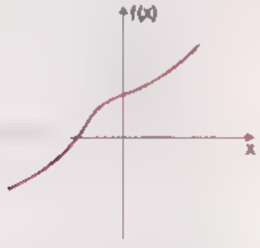
35.1

35.1.0 Introduction

35.1.0

Introduction

We all have an intuitive idea of what is meant by the word *continuous*. For example, we may say that it has rained *continuously* all day, when we mean that it has rained throughout the day without a break. In the same way, we have no difficulty in deciding that one of the two curves shown below is *continuous*, whereas the other is not.



Continuous curve

Discontinuous curve

A definition of *continuity* has been given in *Unit 7, Sequences and Limits I*; in this text we shall generalize this definition and reformulate it in terms of wider applicability.

In making our ideas of continuity more precise, we often use the concept of "nearness". Thus, if we want to determine whether or not a function is continuous at some given point P in its domain, we carry out a formal investigation of what happens *near* the point P. But the word *near* represents only an intuitive concept, and we can see this by asking a few very simple questions:

Is Coventry *near* Birmingham?
Is 1.2067 *near* 1.2063?
Is $x + \varepsilon$ *near* x?

Clearly, a motorist and a hiker might give different answers to the first question. The answer to the second question will depend on the kind of accuracy required in any particular situation in which the numbers arise. The answer to the third question will depend on the general mathematical context and the size of ε.

35.1.1 Distance

35 1 1

Discussion

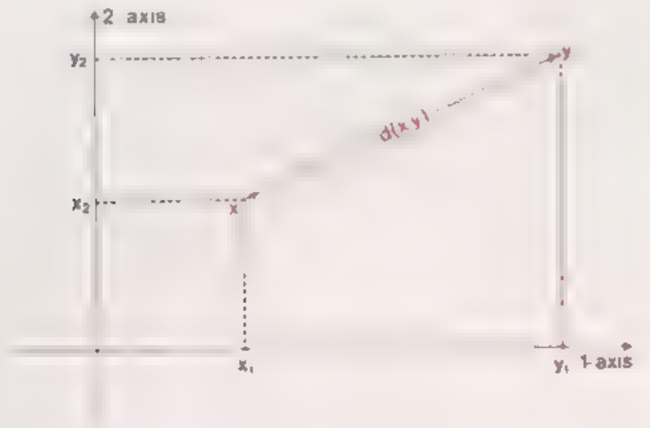
If we want to be precise in formulating and answering questions such as those above, then we need to have some standard of "nearness": one way of achieving this is to define the concept of *distance*. If x and y are two points of R, then we are already familiar with the mathematical definition by which the distance between x and y (thought of as points on the number line) is given by

$$d(x, y) = |x - y|.$$

In the plane, if x and y are two points of R^2, then

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

where (x_1, x_2) and (y_1, y_2) are the co-ordinates of x and y respectively. (We take the positive square root.)



We can generalize the expression for distance to any dimension n. Thus, if x and y are points of R^n, we have

Definition 1

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2},$$

where $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_n)$ are the co-ordinates of x and y respectively. (This expression is simply a generalization based on Pythagoras' theorem.)

Notice that *distance* defines a mapping of $R^n \times R^n \longrightarrow R$, since for any two points $x, y \in R^n$ we have

$$d((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) \longmapsto d(x, y),$$

and $d(x, y) \in R$

Exercise 1

Exercise 1
(2 minutes)

Find the distance $d(x, y)$ in the following cases:

(i) $x, y \in R^2$, where

x is $(-1, 1)$ and y is $(2, -5)$.

(ii) $x, y \in R^4$, where

x is $(2, -1, 0, 3)$ and y is $(-1, 0, 1, -2)$. ■

Solution 1

Solution 1

(i) $d(x, y) = \sqrt{(-1 - 2)^2 + (1 - (-5))^2}$

$= \sqrt{9 + 36} = 3\sqrt{5}$

(ii) $d(x, y) = \sqrt{(2 - (-1))^2 + (-1 - 0)^2 + (0 - 1)^2 + (3 - (-2))^2}$

$= \sqrt{9 + 1 + 1 + 25}$

$= 6.$ ■

Discussion

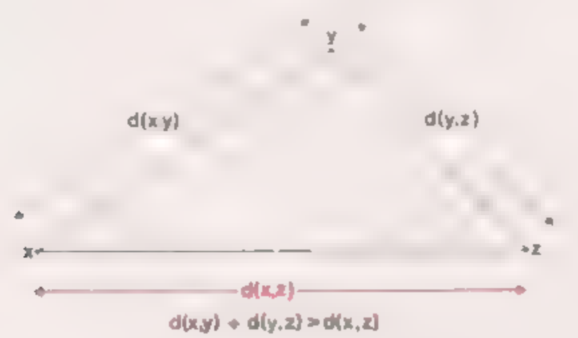
As so often happens in the study of mathematics, what really interests us is not so much the actual formula by which some quantity can be calculated, but the general properties which such entities possess. The distance $d(x, y)$ is a real number which satisfies the following conditions:

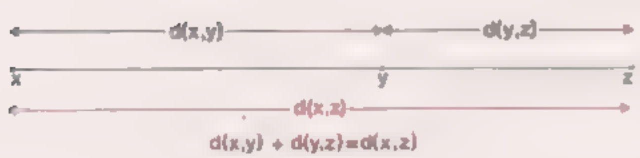
(i) $d(x, y) \geqslant 0$
(ii) $d(x, y) = 0 \Leftrightarrow x = y$ $\qquad (x, y, z \in R^n)$
(iii) $d(x, y) = d(y, x)$
(iv) $d(x, y) + d(y, z) \geqslant d(x, z)$

The first condition states that the distance $d(x, y)$ is always a non-negative real number. The second states that the distance from any point to itself is zero. The third states that distance is symmetric. The fourth is the well-known *triangle inequality*, so called because when x, y, z are taken as the three vertices of a plane triangle, the condition is derived from the theorem that the sum of the lengths of any two sides of the triangle is greater than the length of the third side. The special case

$$d(x, y) + d(y, z) = d(x, z)$$

arises only when x, y, z all lie on the same straight line.





(We have met the *triangle inequality* before:
in the context of geometric vectors (*Unit 22*, section 22.1.3);
in the context of complex numbers (*Unit 27*, section 27.4.3);
in the context of norms of vectors (*Unit 28*, section 28.2.2).)

35.1.2 Continuity

35.1.2

Main Text

In *Unit 7, Sequences and Limits I*, we gave the following definition of a continuous function $f: R \longrightarrow R$:

> If f is a real function and a is an element of its domain, then f is continuous at a if $\lim_{x \to a} f(x)$ exists and is equal to $f(a)$.

Looking at the definition of $\lim_{x \to a} f(x)$ given in *Unit 7**, we find that this limit is equal to $f(a)$ if, given any positive number ε, however small, there is a positive number δ such that the image under f of the non-empty set $\{x : x \in R, 0 < |x - a| < \delta\}$ is a subset of $]f(a) - \varepsilon, f(a) + \varepsilon[$.

We are now able to generalize the notion of a continuous function. All we need to do is to replace any explicit or implicit reference to the modulus, which is appropriate while we are working entirely in R, by the appropriate distance function d, which is its generalization when we work in R^n. Thus $|x - a| < \delta$ becomes

"the distance between a and x is less than δ",

i.e.

$$d(a, x) < \delta,$$

and $f(x) \in]f(a) - \varepsilon, f(a) + \varepsilon[$ holds if

"the distance between $f(a)$ and $f(x)$ is less than ε",

i.e.

$$d(f(a), f(x)) < \varepsilon.$$

So we can now define a continuous function from a subset of R^m to a subset of R^n as follows.

Let

$$f: X \longrightarrow Y,$$

where $X \subseteq R^m$ and $Y \subseteq R^n$.

At any point $a \in X$, we say that f is **continuous at a** if, given any $\varepsilon > 0$, there exists $\delta > 0$ such that, if $x \in X$ and

Definition 1

$$d_m(a, x) < \delta,$$

then

$$d_n(f(a), f(x)) < \varepsilon.$$

(We use d_m and d_n to distinguish distances in R^m from those in R^n.)

Example 1

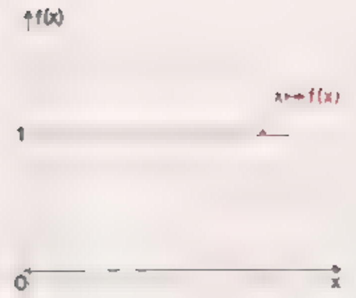
Example 1

Consider the function $f: R \longrightarrow R$ defined by

$$f: x \longmapsto 0 \qquad (x \leqslant 0)$$

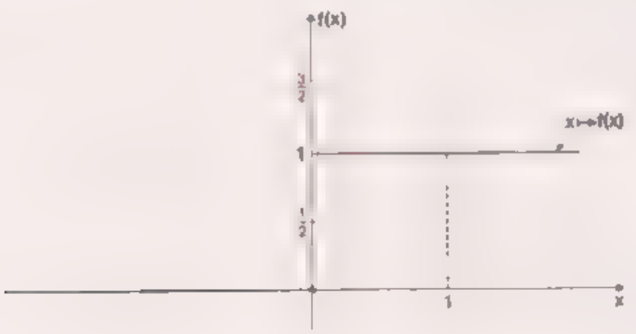
$$f: x \longmapsto 1 \qquad (x > 0)$$

* We have slightly modified the definition given in *Unit 7*. We have replaced $|x - a| \leqslant \delta$ by $|x - a| < \delta$ and replaced $[f(a) - \varepsilon, f(a) + \varepsilon]$ by $]f(a) - \varepsilon, f(a) + \varepsilon[$. We have "toughened up" a bit by excluding the end-points of the intervals in the domain and image set.



Intuitively, we see that f is continuous everywhere except at 0.

Let us look at the point 1. Then the image, $f(1)$, is 1 also. Now choose some arbitrary value for ε, say $\varepsilon = \frac{1}{2}$. In the codomain, we are considering, therefore, the interval $]1 - \frac{1}{2}, 1 + \frac{1}{2}[$, i.e. $]\frac{1}{2}, \frac{3}{2}[$.

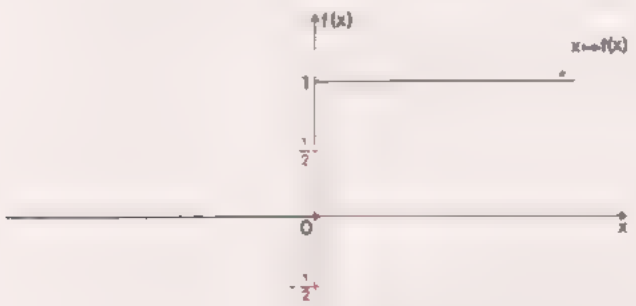


We want to find $\delta > 0$ such that when the distance of x from 1 is less than δ, the distance of $f(x)$ from $f(1)$ is less than ε. That is, we have to find some number $\delta > 0$ giving an interval $]1 - \delta, 1 + \delta[$ whose image lies within $]\frac{1}{2}, \frac{3}{2}[$. Clearly, any $\delta < 1$ will satisfy the required condition, since all points of the domain lying within a distance less than unity from 1 all map to 1 in the codomain, and their images are therefore in $]\frac{1}{2}, \frac{3}{2}[$. In fact $\delta < 1$ is suitable for *any* other choice of ε. A similar argument applies if we choose any other point in the domain, *except* 0.

We now consider the point 0. Since we have

$$f : x \longmapsto 0 \qquad (x \leqslant 0),$$

the image of 0, $f(0)$, is 0. Now choose $\varepsilon = \frac{1}{2}$ once again. This time we consider the interval $]-\frac{1}{2}, \frac{1}{2}[$ in the codomain.



Now, no matter what $\delta > 0$ we choose, we are bound to have some point within distance δ of 0 in the domain whose image is 1, since

$$f: x \longmapsto 1 \qquad (x > 0).$$

So some point within distance δ of 0 is bound to have image 1, no matter how small we choose δ, and 1 lies *outside* the interval $]-\frac{1}{2}, \frac{1}{2}[$. So the function is *not* continuous at 0, and this confirms the intuitive idea we had when we looked at the graph of the function. ■

A function is said to be continuous if it is continuous *at every point* of its domain. (Notice the distinction between *continuous at a point* and *continuous*: the latter means continuous everywhere.) Definition 2

We shall return to the study of continuity later. The definition we have at the moment will suffice for our immediate purpose, which is to discover just what topology is all about

35.2 GEOMETRIES AND INVARIANTS

35.2

35.2.0 Introduction

35.2.0

Introduction

We can regard topology as, in essence, a very fundamental kind of geometry. We shall follow Klein's approach and regard each type of geometry as distinguished by a particular class of transformations.

Before introducing and discussing the special kind of transformation with which we are specifically concerned in the study of topology, we shall consider a few transformations $R^2 \longrightarrow R^2$, and, in particular, the properties *which are preserved under these transformations.*

In *Unit 30, Groups I*, we considered sets of mappings (symmetry operations) which left given figures invariant. We found that the set of all symmetry operations which leave a given figure invariant form a group under composition of functions. In section 35.2 we consider a few geometries; any particular geometry is the study of those properties of geometric objects which are invariant under a set of permitted transformations. In each case the set of one-one transformations which leave a given property invariant necessarily form a group under composition. In each geometry, we say that two figures are *equivalent* if one figure can be obtained from the other by the type of transformation being considered.

We shall see a kind of pattern emerging. Each geometry includes those which precede it. As we increase the variety of permitted transformations, so the number of invariants decreases, and more and more objects become equivalent.

35.2.1 Isometries

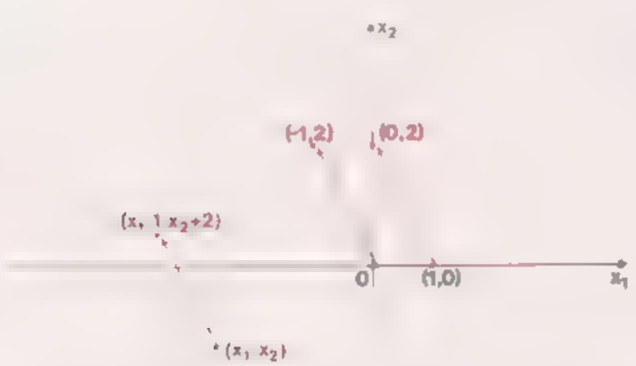
35.2.1

Discussion

We shall consider the transformation $f: R^2 \longrightarrow R^2$ given by

$$f:(x_1, x_2) \longmapsto (x_1 - a, x_2 - b).$$

(Notice that we are using x_1 and x_2 as the co-ordinates in R^2 of some general point x.) Suppose that $a = 1$ and $b = -2$. Then each point of R^2, (x_1, x_2), is *translated* to a new point in R^2, $(x_1 - 1, x_2 + 2)$.



Let us see what has happened to the *distance* between points under this transformation. (Intuitively, we know that distance does not change under a translation, but now that we are trying to obtain precision, we check the algebra.)

Suppose we have two points $(0, 5)$ and $(-2, 1)$. The distance between them is

$$\sqrt{(0 + 2)^2 + (5 - 1)^2} = \sqrt{20}.$$

If we now *translate* these points with $a = 1$, $b = -2$, we obtain the new points $(-1, 7)$ and $(-3, 3)$ respectively. The distance between them is

$$\sqrt{(-1 + 3)^2 + (7 - 3)^2} = \sqrt{20},$$

the same value as before. In fact, it is easy to show that the distance between any two points will always be preserved by the general transformation

$$f:(x_1, x_2) \longmapsto (x_1 - a, x_2 - b).$$

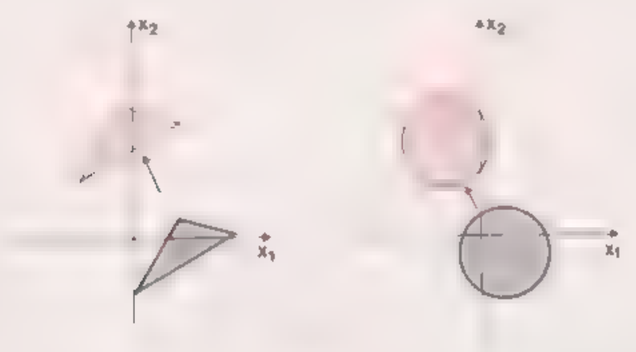
To do this, we consider two points, x and y, of R^2, with co-ordinates (x_1, x_2) and (y_1, y_2) respectively. We have

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2},$$

and

$$\begin{aligned} d(f(x), f(y)) &= \sqrt{((x_1 - a) - (y_1 - a))^2 + ((x_2 - b) - (y_2 - b))^2} \\ &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \\ &= d(x, y). \end{aligned}$$

This means that, if any subset of R^2, such as a triangle, a circle, etc., is *translated*, the exact shape and size will be preserved: only the position will be changed.



Next, we consider the transformation $R^2 \longrightarrow R^2$ given by

$$f:(x_1, x_2) \longmapsto (x_1 \cos\psi - x_2 \sin\psi, x_1 \sin\psi + x_2 \cos\psi).$$

We can write this in matrix notation:

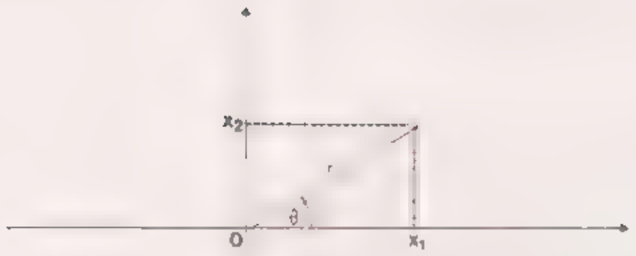
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \longmapsto \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

We know that this transformation is a *rotation* about the origin through an angle ψ anti-clockwise. (See *Unit 23*, section 23.2.4.)

Is distance preserved under this transformation? We can answer this more easily if we express the distance between two points of R^2 in terms of their polar co-ordinates. We change from Cartesian to polar co-ordinates using the usual convention that

$$x_1 = r\cos\theta,$$

$$x_2 = r\sin\theta.$$



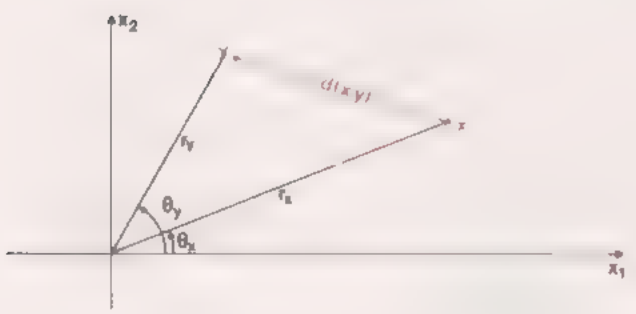
If $x = (r_x, \theta_x)$ and $y = (r_y, \theta_y)$ are polar co-ordinates of two points x and y, then their Cartesian co-ordinates are

$$x = (r_x \cos\theta_x, r_x \sin\theta_x)$$
$$y = (r_y \cos\theta_y, r_y \sin\theta_y).$$

Hence we have

$$\begin{aligned} d(x, y) &= \sqrt{(r_x \cos\theta_x - r_y \cos\theta_y)^2 + (r_x \sin\theta_x - r_y \sin\theta_y)^2} \\ &= \sqrt{\begin{aligned} &r_x^2 \cos^2\theta_x - 2r_x r_y \cos\theta_x \cos\theta_y + r_y^2 \cos^2\theta_y \\ &\quad + r_x^2 \sin^2\theta_x - 2r_x r_y \sin\theta_x \sin\theta_y + r_y^2 \sin^2\theta_y \end{aligned}} \\ &= \sqrt{r_x^2 + r_y^2 - 2r_x r_y \cos(\theta_x - \theta_y)} \end{aligned}$$

(See RB10)



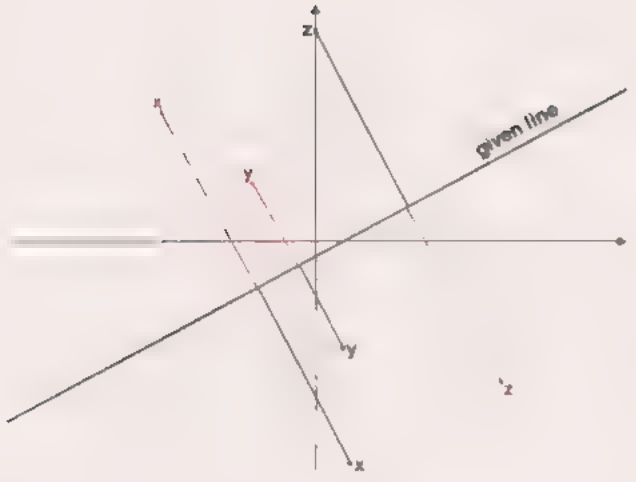
So, if x and y are rotated through ψ about the origin to new points $(r_x, \theta_x + \psi)$, $(r_y, \theta_y + \psi)$, the distance between these points is then

$$\sqrt{r_x^2 + r_y^2 - 2r_x r_y \cos((\theta_x + \psi) - (\theta_y + \psi))} = d(x, y).$$

Now we may ask if the same preservation of distance holds good for rotations about points other than the origin. We can see that this must be so, since we can perform a translation so that our origin moves to the new centre of rotation; the rotation now becomes a rotation about the origin, preserving distance; then we can perform a translation to move the origin back to its original position, again preserving distance. (Intuitively, distance does not depend on the choice of co-ordinates, so rotating about the origin or any other point must both have the same effect on distance.)

So we have *translation* and *rotation* as two distance-preserving transformations.

Another simple transformation of R^2 is the one which *reflects* each point in a given straight line. For example, the points x, y, z in the following diagram are mapped to the points x', y', z'. We can see intuitively from the diagram that all distances between pairs of points in R^2 are preserved by this transformation.



Exercise 1

Exercise 1
(4 minutes)

Show that distance in R^2 is preserved by the transformation which reflects each point $x \in R^2$ in any given straight line. ■

Discussion

The three transformations $R^2 \longmapsto R^2$, *translation*, *rotation* and *reflection*, all have this important property of preserving the distance between any two points. That is, if $x \longmapsto x'$, $y \longmapsto y'$ under one of these transformations, then

$$d(x, y) = d(x', y').$$

Definition 1

We therefore say that *distance* is *invariant* under these transformations. Elementary geometry embraces a study of the properties which are invariant under these so-called rigid transformations or isometries. Two figures are *equivalent* in elementary geometry, i.e. *congruent*, if one may be obtained from the other by means of one or more isometries. (The relation "is congruent to" is an equivalence relation.) It can be shown that any isometry can be expressed as the composition of the three particular types of transformation we have discussed. In fact, there is an even stronger result: any isometry is the composition of at most three reflections. (For proofs of these statements see, for example, H. Liebeck, *Algebra for Scientists and Engineers* (John Wiley, 1969).)

We have already discussed translations, rotations and reflections in *Unit 30, Groups I*. The identity isometry is "do nothing". The inverse of a translation is a translation; the inverse of a rotation through θ is a rotation through $-\theta$; a reflection is its own inverse. We know that certain sets of isometries form groups under the composition of functions. In particular, the set of all isometries of the plane forms a group of infinite order.

Solution 1

Solution 1

First consider reflections in the 1-axis. We have

$$(x_1, x_2) \longmapsto (x_1, -x_2)$$
$$(y_1, y_2) \longmapsto (y_1, -y_2).$$

Hence

$$\begin{aligned} d((x_1, -x_2), (y_1, -y_2)) &= \sqrt{(x_1 - y_1)^2 + (-x_2 - (-y_2))^2} \\ &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2} \\ &= d((x_1, x_2), (y_1, y_2)) \end{aligned}$$

i.e.

distance is preserved.

But a reflection in *any* line which intersects the x-axis is equivalent to a rotation (to send the given line to the x-axis) followed by a reflection in the x-axis, followed by the reverse of the original rotation. If the line is parallel to the x-axis, then we require a translation instead of a rotation. Hence distance is still preserved. ■

35.2.2 Similarities

35.2.2

Discussion

We consider now a different kind of transformation $f: R^2 \longrightarrow R^2$, one that multiplies all distances by a fixed factor $m > 0$. This means that if $x, y \in R^2$, then

$$d(f(x), f(y)) = m \times d(x, y).$$

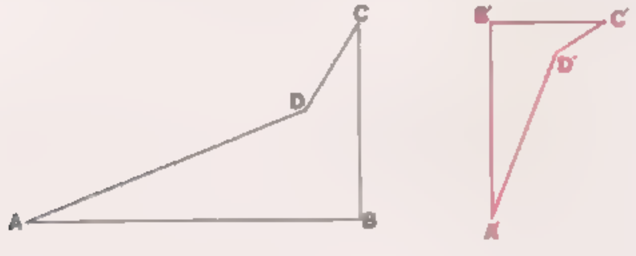
Such transformations are called similarities.

Definition 1

When $m > 1$, we call the transformation an *expansion*. When $m < 1$, we call the transformation a *contraction*. When $m = 1$, the transformation is an isometry.

Now, by definition, distance is *not* invariant under expansion and contraction. We can see, however, that a number of invariants do exist. Straight lines are preserved as straight lines, angles between lines are preserved, and "shape" is preserved.

The study of properties which are preserved under expansion and contraction, together with the isometries, is called *similarity geometry*. Two figures are *equivalent* in this geometry, i.e. *similar*, if one can be obtained from the other by means of one or more similarities. (The relation "is similar to" is an equivalence relation.) In such a geometry, *equivalence* is no longer the same as *congruence*. Congruent shapes are indeed still equivalent, but all similar triangles, for example, are also equivalent. Thus the two shapes shown below are equivalent.



In similarity geometry all circles are equivalent, all cubes are equivalent, all spheres are equivalent, and so on.

The inverse of an expansion is a contraction and vice versa. Certain sets of similarities form groups under composition. In particular, the set of all similarities of the plane forms a group of infinite order.

Exercise 1

Exercise 1
(2 minutes)

Which of the following are invariants of similarity geometry?

(i) ratio of distances;
(ii) area;
(iii) regularity of polygons;
(iv) parabolic shape;
(v) perimeter length;
(vi) parallel property. ■

Solution 1 Solution 1

(i) Invariant.
(ii) Not invariant. Area is a product of distances, and distance is not invariant.
(iii) Invariant.
(iv) Invariant.
(v) Not invariant. Perimeter is a sum of distances.
(vi) Invariant. ■

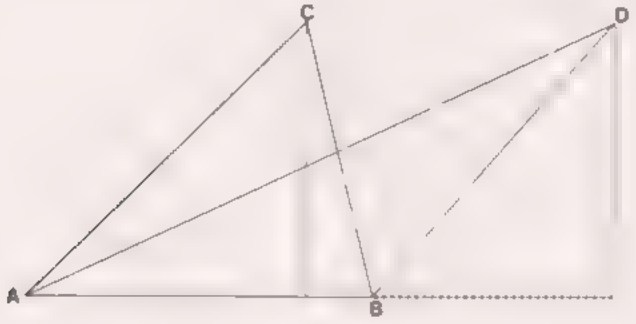
35.2.3 Other Geometries

35.2.3 Discussion

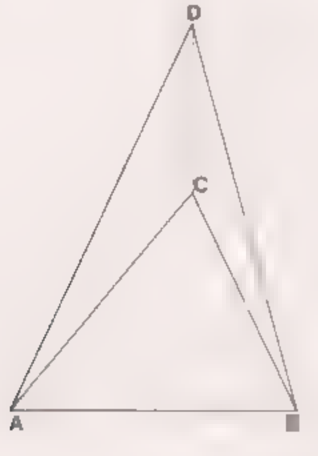
There are other geometries in which further transformations, other than similarities are permitted. If in R^2 we allow the transformation

$$(x_1, x_2) \longmapsto (ax_1 + bx_2 + l, cx_1 + dx_2 + m),$$

where $a, b, c, d, l, m \in R$ and $ad \neq bc$, then we find that distance, area, angle, and even shape are not preserved. Squares can be mapped to parallelograms, circles can be mapped to ellipses. So all kinds of different shapes can be equivalent in the new geometry; this geometry is called *affine geometry*. What are the invariants under such a transformation? In fact, there are still a number of very important invariants: points of a line still divide line-segments in the same ratio, finite configurations remain finite, parallel lines remain parallel, and so on. A transformation which is permitted in affine geometry is that known as a *shear*, which maps a triangle ABC to another triangle ABD, having the same base AB and the same height.



If the triangles have the same base but different heights, then the transformation is called a *strain*.



In this geometry, all triangles are equivalent.

(Another geometry, which goes (as it were) one step further, is *projective geometry*. In this geometry it is usual to adopt a special type of co-ordinate system which we shall not discuss. So we shall not give any general expression for the projective transformation in R^2. (Roughly, we can think of it as a perspective projection of a figure from some point outside it, such as occurs when a photograph is taken.) We lose the invariants of parallelism, of the ratio in which points divide line-segments, and of the special property of the finite. A square can be mapped, for example, to any quadrilateral. There are still obvious invariants left however: a straight line remains a straight line, and, in particular, cross-ratio is preserved; that is, if A, B, C and D are distinct points in a straight line, then the ratio of lengths

$$\frac{CA}{CB} \cdot \frac{DA}{DB}$$

is an invariant.)

35.2.4 Topology

35.2.4

Main Text

Definition 1

In order to arrive at *topology*, we extend our permitted transformations and allow any one-one continuous transformation whose inverse is also continuous; such a transformation is said to be bi-continuous. We already have a formal definition of a continuous transformation of R^n, and the requirement that the transformation be one-one means that each point will be mapped to a unique image point which will itself have the original point as its unique image under the inverse transformation. We now therefore permit the so-called *elastic* deformations of stretching, bending, twisting, and so on. We are not, however, allowed to tear, or to join separate bits together, but we may cut and tie knots as long as we exactly repair the cut afterwards. In other words, *adjacent* points must remain adjacent, i.e. *neighbourhoods* must not be disturbed. (It doesn't matter what we do during the process of transformation, so long as the end product conforms to our rules. Putting it another way, the mechanics of the transformation do not bother us; we are only interested in the initial and final positions.)

Definition 2

Definition 3

Topology is the study of the properties of structures which are invariant under these bi-continuous transformations. Such transformations are called topological transformations. It might seem, at first sight, that the invariants remaining must now be so few that almost all structures will be equivalent. This is by no means the case. Topological invariants are, by their very nature, more fundamental than those of the geometries which we have discussed, and for this reason they are of very much wider applicability in mathematics. We shall finish this section by giving two examples of topological invariants in our usual geometric space.

Consider, for example, what happens if we draw any continuous closed curve C on the surface of a sphere. Intuitively, we can see that the curve will separate the surface of the sphere into two distinct parts. To get from a point in one part of the surface to a point in the other part, it is necessary to cross C.

A sphere can be transformed by a topological transformation into a cube, and hence a sphere and a cube are topologically equivalent. If we map the sphere with its closed curve C by a topological transformation on to the cube, then the surface of the new object will still be separated into two distinct parts by the continuous closed curve C' which is the image of the original curve C. So the particular property we have just described is a topological invariant of the transformation.

Consider now an object known as a *torus*, which we may think of as a ring-doughnut or an inner tube. The curves C_1 and C_2 are each continuous closed curves on the surface of a torus.

The curve C_1 has exactly the same property of separating the surface into two parts as the curve C on the sphere, but the curve C_2 does not have this property. It is possible to travel from a point on one "side" of C_2 to a point on the other "side", without crossing C_2, by travelling all the way round the torus. It is not possible to draw a curve like C_2 on the surface of a sphere. The sphere and the torus are *not* topologically equivalent, for if they were, the transformation which would map the torus to the sphere would map C_2 on the torus to a curve on the sphere with the same properties — and no such curve exists on the sphere.

Exercise 1

Exercise 1
(5 minutes)

(i) In what other way can a continuous closed curve be drawn on the surface of a torus without separating the surface?
(ii) What is the largest number of non-intersecting continuous closed curves which can be drawn on the surface of a torus without separating it?
(iii) If the torus has two holes (as shown below), would your answer to (ii) be the same?

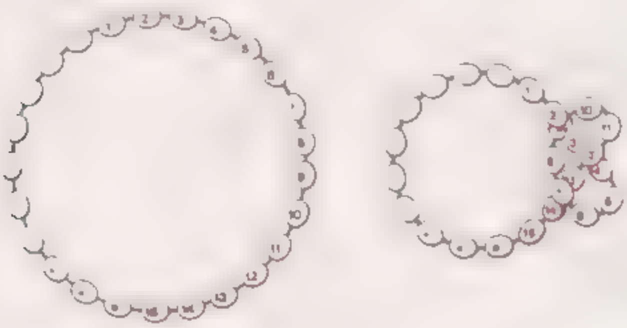
Discussion

The number of closed curves which can be drawn on any particular surface without separating that surface is a *topological invariant*, that is to say, it remains the same under all topological transformations of the surface.

Suppose now that we take a circle, cut it at some point, open it out and tie a knot, and then rejoin the two ends exactly as they were before we cut them. The process is represented by stages (a), (b), (c) in the diagram.

(a) (b) (c)

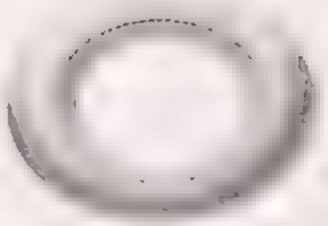
This is still a one-one continuous transformation, and an important property which is preserved is the *order* of points on the circle. We can see this is we imagine beads on a thread numbered 1, 2, 3, . . . , n.



Although we finish up with a knot, after we have cut, knotted and repaired the cut, the actual order of the beads continues round the closed thread in exactly the same way as on the original circle.

Solution 1

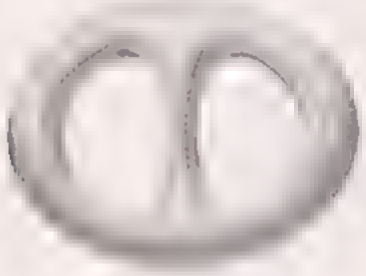
(i)



The dotted line shows a curve with the required property.

(ii) Only one such curve may be drawn without separating the torus.

(iii) The answer would not be the same. Two closed, non-intersecting curves may then be drawn without separating the surface.



■

35.3 OPEN SETS AND TOPOLOGICAL SPACES

35.3

35.3.0 Introduction

35.3.0

Introduction

What we have so far is known as *metric topology*, because we have used the *metric* or distance function $d(x, y)$ in our definition of a continuous transformation.

Since distance ceased to be an invariant as soon as we moved from isometry geometry to similarity geometry, it is certainly not one of the invariants of topology.

Because topology is so fundamental, it is desirable — in fact ultimately essential — that the concept of a continuous transformation should be re-defined in terms which do not include the notion of distance.

We cannot do this, however, without further examination of the concept of "nearness". To begin with, we must continue to talk in terms of distance (since the metric topology is all that is available to us at this stage), but we shall attempt to formulate non-metric concepts as much as possible.

35.3.1 Neighbourhoods

35.3.1

Main Text

For any point $x \in R^n$, a set of points *near* x may be defined to be the set of all points $y \in R^n$ for which

$$d(x, y) < r,$$

for some fixed real number $r > 0$. These points form what is called a *neighbourhood* of x in R^n, consisting of just those points of R^n whose distance from x is less than r. This neighbourhood is denoted by

$$N(x, r, R^n).$$

For example, in two dimensions we might have the following:

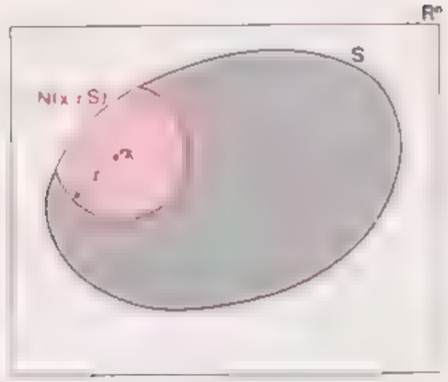
In this case, the neighbourhood is just the set of points which lie inside a circle with centre x and radius r.

More generally, for any set S (contained in R^n, so that the metric is defined) and for any $x \in S$ and $r > 0$, a *neighbourhood* of x in S is the set of all points y in S for which

Definition 1

$$d(x, y) < r.$$

This neighbourhood is denoted by $N(x, r, S)$. Notation 1



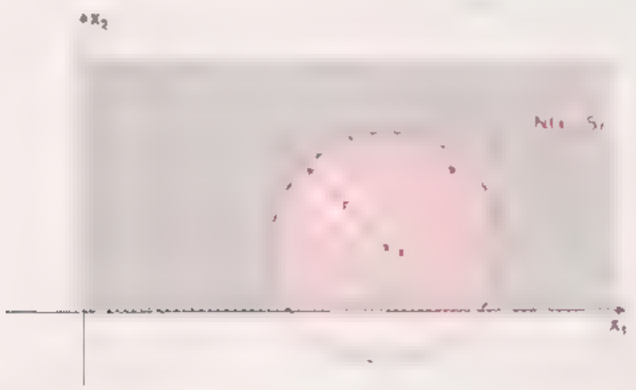
Notice that x belongs to each of its neighbourhoods.

In the remainder of this text, we shall frequently use S to denote the "space" in which we are working at the time; points outside S do not concern us — to all intents and purposes they do not exist (for us). The only reason we write $S \subseteq R^n$ is to tell us what the metric (distance function) is: we need to know this in order to determine the neighbourhoods of points in S.

As we shall see later, a "space" is just a set in which neighbourhoods are defined — but first we must pursue the concept of a neighbourhood further.

Example 1 Example 1

Again in two dimensions, suppose that S is the subset of R^2 comprising all the points having both co-ordinates positive.



This time the neighbourhood of x is not a complete disc the part of the disc on or below the horizontal axis is excluded. ■

Exercise 1 Exercise 1 (2 minutes)

Describe the following neighbourhoods.

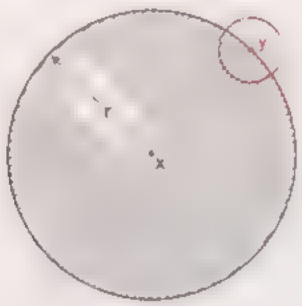
(i) $N(x, r, R)$
(ii) $N(x, r, R^3)$. ■

35.3.2 Boundary

35.3.2

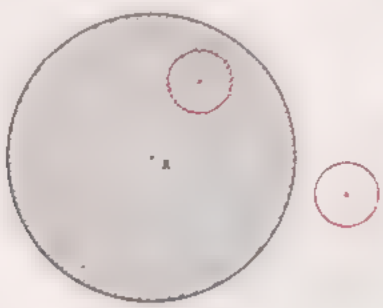
Discussion

Notice that our definition of a neighbourhood has been framed so as to exclude all points at distance exactly r from x; that is to say, the boundary is not a part of the neighbourhood. Let us see if we can define precisely what we mean by *boundary*.



We show a circular disc in R^2, with centre x and radius r, and a point y on the circumference of the circle itself. We also show a typical neighbourhood of y.

With the disc shaded, it is easy to see that any neighbourhood of y in R^2 will have both a shaded and an unshaded region. In other words, it must include some points in $N(x, r, R^2)$ and also some points not in $N(x, r, R^2)$.



The same is not true of points in R^2 which are not on the boundary, as the above diagram shows. Some points have (some) neighbourhoods all shaded, and some points have (some) neighbourhoods wholly unshaded. In fact, any point at distance $r - s$ or $r + s$ from x has all its neighbourhoods of radius s or less either entirely shaded or entirely unshaded.

Main Text

The boundary is thus characterized as the set of points which have every one of their neighbourhoods *partly inside* and *partly outside* the disc.

Encouraged by this discovery, we generalize the concept of boundary, and define the boundary for any set in R^n.

Definition 1

Definition 2

If $X \subseteq S \subseteq R^n$, then a point $p \in S$ is said to be a boundary point of X in S if *every* neighbourhood of p in S includes at least one point of X and at least one point (in S) not in X. The boundary of X in S is that subset of R^n consisting of all the boundary points of X.

For example, the boundary of $[a, b]$ or $]a, b[$ in R is $\{a, b\}$.

You may wonder what we have achieved by this definition of a perfectly obvious concept. We have defined *boundary* in terms of *neighbourhood*,

(continued on page 22)

Solution 35.3.1.1

Solution 35.3.1.1

(i) We require the set of real numbers which differ from x by less than r. Hence

$$N(x, r, R) = \{y : x - r < y < x + r\} =]x - r, x + r[.$$

(ii) Here $N(x, r, R^3)$ is the interior of the sphere, centre x and radius r. ■

(continued from page 21)

and we shall go on to define one or two other concepts in terms of these two. Eventually we shall propose a more general definition of *continuity*. Remember that it is our object to define continuity in entirely non-metric terms. Neighbourhood does however contain the idea of distance in its definition, so it may look as if we are not progressing. This is not the case, however: we shall eventually replace neighbourhood itself, thus obtaining an entirely metric-free definition of continuity.

35.3.3 Open and Closed Sets

35.3.3

In the case of a neighbourhood as defined in section 35.3.1, the boundary is specifically excluded. This prompts the following definitions.

Main Text

Any subset of $S \subseteq R^n$ which includes *none* of its boundary points is said to be an open set in S.

Definition 1

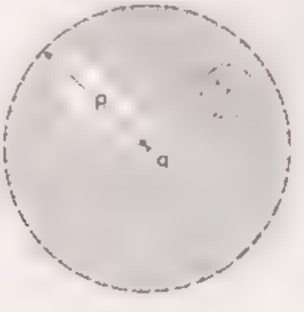
For example, the open interval $]a, b[$ is an open set in R.

Any subset of $S \subseteq R^n$ which includes *all* its boundary points is said to be a closed set in S.

Definition 2

For example, the closed interval $[a, b]$ is a closed set in R.

Open sets have a very significant property: since no point of an open set can be a boundary point, it follows that every point x of an open set X in S must have some neighbourhood which does *not* include both points in X and points not in X (unlike a neighbourhood of a boundary point). Since x is certainly in every one of its neighbourhoods, we deduce that it must have at least one neighbourhood lying entirely in X.



As an example, we consider the open disc in R^2, with centre q and radius ρ. For any x inside the disc, and so at some distance $\rho - d$ from the centre q, the neighbourhood $N(x, r, R^2)$ lies entirely inside the disc provided $r \leqslant d$.

By definition, neighbourhoods are always open — but remember that openness is a property relative to the containing set S. We often take S to be R^n itself, but sometimes it is a subset of R^n. The following exercise illustrates this point.

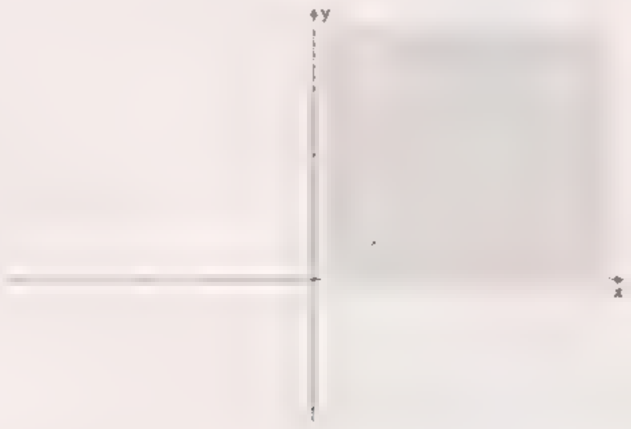
Exercise 1

Exercise 1
(4 minutes)

(i) Show that the neighbourhood $N(x, r, S)$ in Example 35.3.1.1 is an open set in S.
(ii) Show that the same is still true if S includes the positive half of the horizontal axis. ■

Main Text

Many sets include some, but not all, of their boundary points: they are therefore neither open nor closed.



The shaded area above shows the subset of R^2 consisting of all points with co-ordinates (x, y) for which $x \geqslant 0$ and $y > 0$.

The boundary of this subset in R^2 consists of the non-negative half of each axis. Whereas the strictly positive half of the vertical axis belongs to the set, neither the origin nor any part of the horizontal axis is included. Hence the set is *neither* open *nor* closed: it includes some but not all of its boundary points.

Strangely enough, it is possible for a set to be *both* open and closed. Consider the empty set, $\varnothing$. Since $\varnothing$ contains no points at all, it certainly does not contain any boundary points: hence it is open, by definition. But, from the definition of a boundary point, a set without points can have no boundary points (every neighbourhood of a boundary point must include a point of the set, and this set has no points); so the boundary of $\varnothing$ is again the empty set, $\varnothing$ itself Since $\varnothing$ certainly contains $\varnothing$, we must acknowledge that the empty set includes the whole of its boundary; hence $\varnothing$ is closed, by definition. We can find many other examples with these properties. Another set which is both open and closed is the set S itself. Any neighbourhood of a point of S, by definition, contains only points of S, so the boundary of S is $\varnothing$ in S, and S is therefore open. For the same reason, S is also closed; since it has boundary $\varnothing$, S technically contains the set of all its boundary points, the empty set.

Exercise 2

Exercise 2
(4 minutes)

(i) Show that the interval $[0, 1[$ in R, that is, the set of all $x \in R$ for which

$$0 \leqslant x < 1,$$

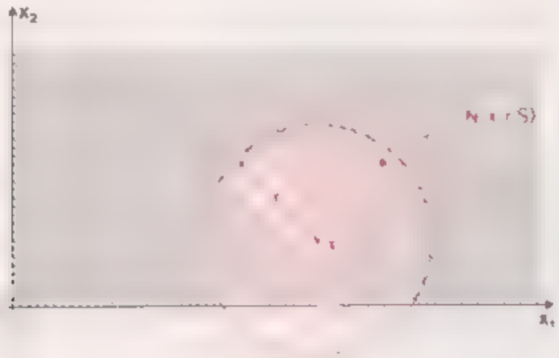
is neither open nor closed in R.
(ii) By considering neighbourhoods, demonstrate that $[0, 1[$ is both open and closed when considered as a subset of itself. ■

(*continued on page 24*)

Solution 1

Solution 1

(i) The boundary of $N(x, r, S)$ consists of that part of the circumference of the circle lying above the horizontal axis. It does *not* include any part of the horizontal axis, since points on the axis are explicitly excluded from S.



Since no point on the circle belongs to $N(x, r, S)$, being at distance r from x, the neighbourhood includes none of its boundary points and so is open, by definition.

(ii) When S does include the horizontal axis, then, apart from the addition of the two points in which the circle intersects the axis, the boundary is still as described above, although $N(x, r, S)$ now includes that part of the horizontal axis lying inside the circle. These points do *not* belong to the boundary — remember, their neighbourhoods are semicircular and do not include points below the axis. Intuition fails in cases like this, partly because intuition leads us to expect a "boundary" to be a closed curve — a property which is not implied by our definition. ■

Solution 2

Solution 2

(i) The boundary of $[0, 1[$ in R consists of the two points 0 and 1, i.e. it is the set $\{0, 1\}$. Of these, 0 belongs to $[0, 1[$, so the interval is not open, and 1 does not belong to $[0, 1[$, so the interval is not closed.

(ii) As a subset of itself, $[0, 1[$ has boundary $\varnothing$, since, for example, the neighbourhoods of 0 with respect to $[0, 1[$ include no negative numbers, but only points of $[0, 1[$.

Since $[0, 1[$ therefore does not include any boundary points, it is open, by definition. Equally, the set $[0, 1[$ includes the empty set — which is its boundary — and hence is closed, by definition. ■

(*continued from page 23*)

Main Text

Although many sets are found to be neither open nor closed, there is no difficulty in constructing either open sets or closed sets from any given set.

For instance, if we remove from any set $X \subseteq S$ all its boundary points in S, we obtain an open set called the interior of X in S.

Definition 3

We can also construct closed sets at will. All we need do is to adjoin to any set X in S all its boundary points in S. By definition, the result must be a closed set in S, which we call the closure of X in S.

Definition 4

When X is an open set in S, *every* point of X is an interior point. Every point of an open set X really is *inside* X — it can be enclosed in a neighbourhood consisting entirely of points of X.

It follows that the union of any number of open sets still has this property — for *any* point of the union must belong to at least one of the sets, and it must have a neighbourhood wholly inside that set and hence in the union. It cannot therefore be a boundary point of the union: hence the union of any number of open sets is again an open set.

With intersections of open sets we must be a little more careful. Consider the following open sets, which are all subsets of R:

$$\{x: -1 < x < 1\}$$

$$\{x: -\tfrac{1}{2} < x < \tfrac{1}{2}\}$$

$$\{x: -\tfrac{1}{3} < x < \tfrac{1}{3}\}$$

$$\cdot$$

$$\left\{x: -\frac{1}{n} < x < \frac{1}{n}\right\}$$

This is an infinite sequence of open sets of R

$$]-1, 1[,\ ]-\tfrac{1}{2}, \tfrac{1}{2}[,\ \ldots,\ ]-\tfrac{1}{n}, \tfrac{1}{n}[,\ \ldots$$

The only point which belongs to each of these sets is the origin, 0. Hence their intersection is the single-point set $\{0\}$. This is certainly not open in R, since $N(0, r, R)$ is not a single point for any $r > 0$; it therefore contains at least one point of R not in $\{0\}$, so $\{0\}$ is not open in R.

Exercise 3

Exercise 3
(5 minutes)

Show that $\{0\}$ is a closed set in R. ■

Discussion

If, however, we restrict our consideration to a *finite* number of intersections of open sets, we find that the intersection of a finite number of open sets is an open set. We shall not prove this result, but you might like to try. Notice that the intersection of two non-overlapping open sets is the empty set, which we already know to be open.

Solution 3

Solution 3

Any neighbourhood of 0 in R contains the point 0 and other points not in $\{0\}$. Hence 0 is a boundary point of $\{0\}$. Clearly there can be no other. So $\{0\}$ includes all its boundary points and hence is closed. ■

35.3.4 Summary

35.3.4

Summary

We now summarize the main points discussed in this section so far:

The distance $d(x, y)$ between two points $x, y \in R^n$ has the following properties:

(i) $d(x, y) \geq 0$
(ii) $d(x, y) = 0 \Leftrightarrow x = y$
(iii) $d(x, y) = d(y, x)$
(iv) $d(x, y) + d(y, z) \geq d(x, z)$.

The neighbourhood $N(x, r, S)$ of x of radius r in S is the set of all points $y \in S$ such that

$$d(x, y) < r,$$

where $x \in S \subseteq R^n$ and $r > 0$.

The boundary in S of a set $X \subseteq S \subseteq R^n$ is the set of all boundary points $p \in S$ such that, for *every* $r > 0$, $N(p, r, S)$ contains at least one point of X and at least one point not in X.

An open set in $S \subseteq R^n$ is any subset of S which contains *none* of its boundary points in S.

A closed set in $S \subseteq R^n$ is any subset of S which contains the *whole* of its boundary in S.

Every point of an open set in S has a neighbourhood wholly contained in S.

The union of any number of open sets is open, and the intersection of a finite number of open sets is open.

The empty set $\varnothing$ and the whole set S are both open and closed.

35.3.5 Continuity

35.3.5

Main Text

You may have noticed that the definitions, apart from that of a neighbourhood, are all expressed in non-metric terms. True, they all depend upon the concept of a neighbourhood, but if we could define *neighbourhood* some other way we should have a structure independent of the concept of distance.

Before trying to do this, we must be sure that the ideas we have developed will do the job required of them: specifically, can we re-define continuity in these terms?

In section 35.1.2 we gave a definition of a continuous function

$$f: X \longrightarrow Y$$

in terms of distance. This can be immediately translated into terms of neighbourhoods

Thus, if $X \subseteq R^m$ and $Y \subseteq R^n$, then a function

$$f: X \longrightarrow Y$$

is said to be continuous at a point $x \in X$ if, given any $\varepsilon > 0$, we can choose $\delta > 0$ such that, whenever

Definition 1

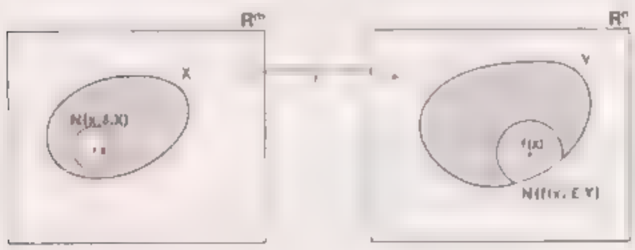
$$y \in N(x, \delta, X), \qquad (y \text{ is "near" } x)$$

then

$$f(y) \in N(f(x), \varepsilon, Y) \qquad (f(y) \text{ is "near" } f(x))$$

i.e.

$$f(N(x, \delta, X)) \subseteq N(f(x), \varepsilon, Y).$$



In other words, $f: X \longrightarrow Y$ is *continuous* at $x \in X$ if for *any* neighbourhood of $f(x)$ in Y there is some neighbourhood of x in X whose image lies within the given neighbourhood of $f(x)$ in Y. In the diagram $N(x, \delta, X)$ and its image $f(N(x, \delta, X))$ are shown in pink; $N(f(x), \varepsilon, Y)$ is the subset of Y indicated in white

Notice that the standard set for "nearness" in the codomain comes first. We demand that, however exacting our standard in the codomain (however small ε), we can meet the requirement with a suitable standard (δ) in the domain.

To see whether a given function $f: X \longrightarrow Y$ is continuous at x we go through the following steps.

(i) Choose an arbitrary neighbourhood of $f(x)$:

$$N(f(x), \varepsilon, Y).$$

(ii) Look for a neighbourhood of x, $N(x, \delta, X)$, such that

$$f(N(x, \delta, X)) \subseteq N(f(x), \varepsilon, Y).$$

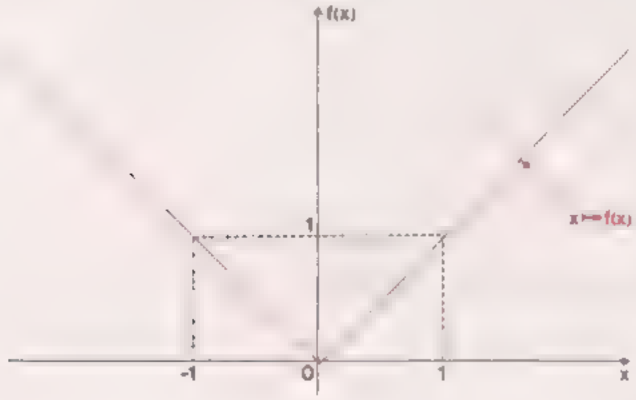
(iii) If such a neighbourhood exists for *any* $\varepsilon > 0$, then the function is continuous at x: otherwise the function is not continuous (i.e. it is discontinuous) at x.

Exercise 1

Exercise 1
(5 minutes)

Carry out a formal investigation of the continuity of the function

$$f: x \longmapsto |x| \qquad (x \in R).$$

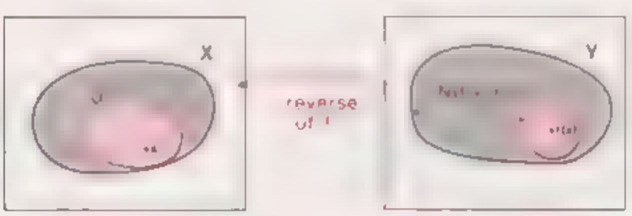


Main Text

Let us now see how the definition of continuity expressed in terms of *neighbourhoods* can be translated into one in terms of *open sets*. The neighbourhood definition tells us that a function $f: X \longrightarrow Y$ is continuous at $x \in X$ if for *any* $\varepsilon > 0$, there is a $\delta > 0$ such that the image under f of (what we may call) the δ-neighbourhood of x lies within the ε-neighbourhood of $f(x)$. So we define f to be continuous if the image under the reverse of f of each open set of the image set (a subset of Y) is an open set of the domain X.

Definition 2

Let's just see why this is so. We start in the image set. Any neighbourhood $N_{f(x)} = N(f(x), \varepsilon, Y)$ of $y = f(x)$ is an open set; so we assume that the neighbourhood $N(f(x), \varepsilon, Y)$ of $f(x)$ is the image of an open set U containing x, in the domain X.



Since U is open in X, we can choose a neighbourhood $N_x = N(x, \delta, X)$ of x entirely within the open set U.

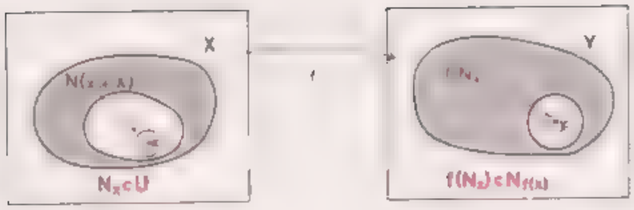
Now

$$f(U) = N_{f(x)}$$

so that

$$N_x \subseteq U \Rightarrow f(N_x) \subseteq N_{f(x)},$$

and so f is continuous at x.



So our new definition of continuity certainly implies the old one. What is not quite so clear is that the reverse also holds true, i.e. that the old definition implies the new one. The two definitions are, however, equivalent.

We shall not prove the equivalence of the two definitions here; we shall, however, look at one or two specific examples. Notice that the open set definition is given in terms of the *reverse mapping* of f.

Example 1

Example 1

First, we look again at $f: R \longrightarrow R$ defined by

$$f: x \longmapsto 0 \qquad (x \leqslant 0),$$

$$f: x \longmapsto 1 \qquad (x > 0).$$

We consider the reverse mapping, g. This gives us

$$g: 0 \longmapsto \{x : x \leqslant 0\},$$

$$g: 1 \longmapsto \{x : x > 0\},$$

that is,

$$g(0) = R_0^-, \qquad g(1) = R^+.$$

Technically, g is defined in the image set $\{0, 1\}$ only, and it is not obvious what to take as the open sets. This can be overcome in many ways. Perhaps the simplest in the context of this course so far is to extend g to have domain R by defining

$$g(x) = \varnothing \qquad (x \in R \text{ and } x \neq 0, 1).$$

Now we can consider the image under g of any open set in R. Suppose that we take the open set $]\frac{1}{2}, \frac{3}{2}[$. The image under g of any point $y \in]\frac{1}{2}, \frac{3}{2}[$ is given by:

$$g(y) = R^+ \qquad (y = 1),$$

$$g(y) = \varnothing \qquad (y \neq 1).$$

So the image of the open set $]\frac{1}{2}, \frac{3}{2}[$ under g is just R^+, but R^+ does not include its boundary '0'. So the image under g of the open set $]\frac{1}{2}, \frac{3}{2}[$ is open in R. This holds good for any open set which does not include zero. So f is continuous at x, provided $x \neq 0$.

If we take $]-\frac{1}{2}, \frac{1}{2}[$, however, as our open set in the domain of g, this does include zero, and for any point $y \in]-\frac{1}{2}, \frac{1}{2}[$ we have:

$$g(y) = R_0^- \qquad (y = 0)$$

$$g(y) = \varnothing \qquad (y \neq 0).$$

(*continued on page 31*)

Solution 1

Solution 1

Consider any $x \in R$. Put $y = f(x)$, i.e. $y = |x|$, so that $y \in R_0^+$.

Then any neighbourhood of $y \in R_0^+$ is of the form either

$]y - \varepsilon, y + \varepsilon[$ with $\varepsilon < y$

or

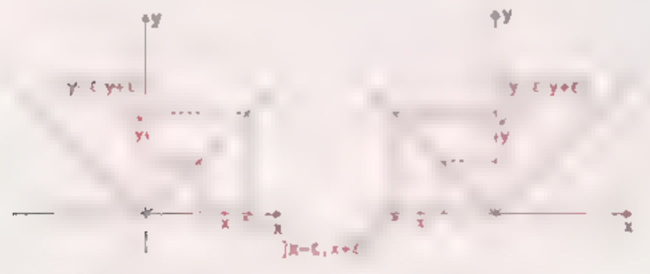
$[0, y + \varepsilon[$ with $\varepsilon \geqslant y$.

But we have

$f(]x - \varepsilon, x + \varepsilon[) =]y - \varepsilon, y + \varepsilon[$ when $\varepsilon < y$

and

$f(]x - \varepsilon, x + \varepsilon[) = [0, y + \varepsilon]$ when $\varepsilon \geqslant y$



the case when ε<y

the case when ε≥y

Consequently every neighbourhood of $y \in R_0^+$ is itself the image under f of a neighbourhood of x in R (and also of a neighbourhood of $-x$ in R). Whence f is continuous at every point of R, i.e. f is continuous.

(Notice that in fact every neighbourhood of $y \in R_0^+$ is the image of a union of *two* neighbourhoods, a neighbourhood of x and a neighbourhood of $-x$. We can tidy things up if we re-phrase our definition of continuity in terms of open sets rather than neighbourhoods.) ■

(continued from page 29)

So the image of the open set $]-\frac{1}{2}, \frac{1}{2}[$ under g is R_0^-. Now R_0^- is closed in R by our definition of a closed set, since it contains its boundary $\{0\}$. Since we have a reverse image of an open set which is not open, the function f is not continuous at 0. ■

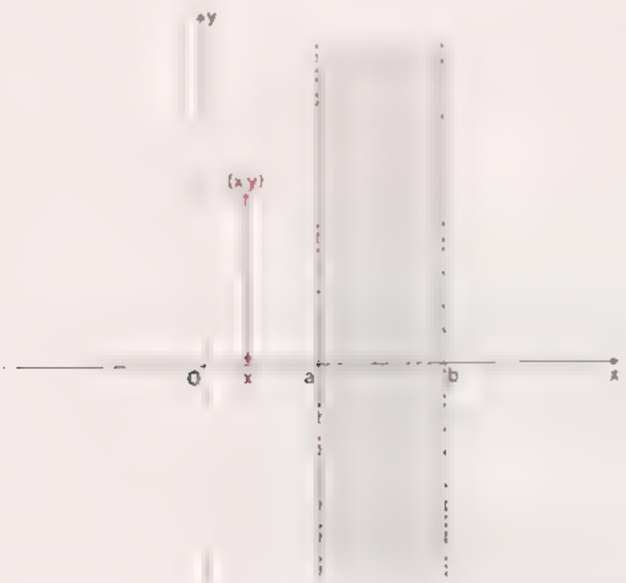
Example 2

Example 2

Consider now $f: R^2 \longmapsto R$ defined by

$$f:(x, y) \longmapsto x \qquad (x, y \in R).$$

(This is simply the "projection mapping" which maps any point given by the ordered pair (x, y) on to its first co-ordinate x.) We shall take the general open interval $]a, b[$ of the codomain R and consider its image under the reverse mapping g.



Whatever the values we choose for a and b, $a < b$, the image under g of the open set $]a, b[$, i.e. $g(]a, b[)$, will always be an infinite vertical strip (as shown) with the edges of the strip excluded. The edges are excluded because they are the images under g of the points a and b, which are not members of the open interval $]a, b[$. So the image under g of $]a, b[$ is always an open set in R^2.

The function

$$f:(x, y) \longmapsto x \qquad (x, y \in R)$$

is therefore a continuous function. ■

Example 3

Example 3

Finally, we look again at the function

$$f: x \longmapsto |x| \qquad (x \in R).$$

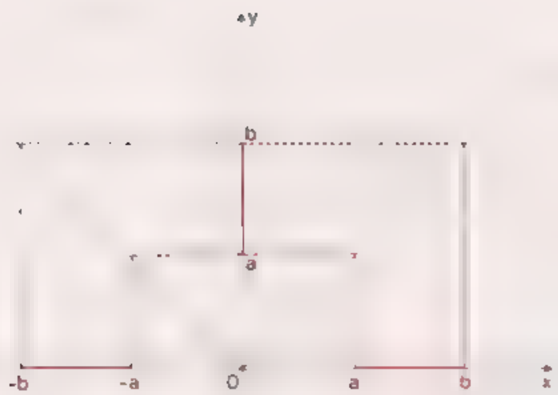
Let $]a, b[$ be some open interval in the codomain R_0^+. The image under g, the reverse of f, of any point $y \in]a, b[$ is given by

$$g(y) = \{-y, y\},$$

and the image under g of the open interval $]a, b[$ is given by

$$g(]a, b[) =]-b, -a[\cup]a, b[.$$

$]-b, -a[$ and $]a, b[$ are both open in R, so their union is open in R.



However we choose a and b, $a < b$, from R_0^+, we shall always obtain an open set in R for $g(]a, b[)$. The intervals $[0, c[$ are also open in R_0^+. The image under g of such an interval is the open interval $]-c, c[$. Consequently every open set in R_0^+ has an inverse image which is open in R. This confirms our previous result, that the modulus function is continuous. ■

Discussion

You may wonder why mathematicians bother to be so precise in the definition of continuous functions, when it might appear that an intuitive approach is likely to give us the same result as a formal application of the definition. The reason is not simply that mathematicians like to have everything "cut and dried". We have pointed out in earlier units that there are functions for which we cannot draw pictorial graphs. Even if we confine our attention to functions $f: R^m \longrightarrow R^n$, we are going to find ourselves in difficulty when m, n are greater than 3. There are also certain so-called "pathological" functions which, although having domain and codomain R, nevertheless behave in very extraordinary and complicated ways, which would defeat any purely geometric and intuitive attempt to determine whether or not they are continuous. But even these cases are only a weak justification of all the trouble we have been to in this unit. The real justification lies in the future: we have succeeded in making our definition of continuity independent of distance, because we now have it in the following form.

A transformation $f: R^m \longrightarrow R^n$ is continuous if the image under the reverse of f of each open set in R^n is an open set in R^m.

This definition will allow us in the end to apply the ideas of continuity to much more general (and possibly simpler) spaces than R^m and R^n, where intuition would hardly apply. And we can be certain that any new results we obtain will not contradict intuition, because our definition includes the definition based on the concept of distance but is more general. Topology has become a vast and important subject in mathematics and we shall have more to say about it in later courses.

We shall conclude this unit by moving away from R^n and R^m. We shall define a general *topological space* and discuss one further example of an old idea, previously based on the concept of distance, which can be redefined in much greater generality.

35.3.6 Topological Spaces

35 3 6

A topological space may be defined as a set S, together with a non-empty collection $\mathscr{C}$ of subsets of S, called open sets, satisfying the following four axioms: Definition 1 Definition 2

T1 The empty set $\varnothing \in \mathscr{C}$.
T2 $S \in \mathscr{C}$
T3 The union of any members of $\mathscr{C}$ is a member of $\mathscr{C}$.
T4 The intersection of any finite number of members of $\mathscr{C}$ is a member of $\mathscr{C}$

The collection $\mathscr{C}$ of subsets is called a topology in (or for) S. Definition 3

This is a very simple definition. Notice that it arises from our experience. We call the subsets of S *open sets* because they have properties similar to those of the open sets considered in the previous sections of this text.

Example 1

Example 1

The set R, together with the collection of subsets defined by all the open intervals $]a, b[$ and all their unions, is a topology It is often called the *usual* topology in R. ■

Example 2

Example 2

We can, however, define other topologies in R. For instance, the smallest possible collection of subsets satisfying these axioms is $\{\varnothing, R\}$. Going to the other extreme, the set of *all* subsets of R is also a topology in R. With this topology every subset of R is called an *open set*. These two topologies can obviously be defined on any set S. ■

Exercise 1
(5 minutes)

Exercise 1

Find all the topologies on a set of three elements. ■

A topological space is more general than a space upon which distance is defined. When we considered continuity, we started by using the concept of distance, but we ended up by defining continuity also in terms of the preservation of open sets under the reverse transformation, so as to make our definition more general. Discussion

We can now adopt this definition of continuity for a function linking two general topological spaces:

If S and T are sets which have topologies defined on them and $f: S \longrightarrow T$, then f is said to be continuousif the image of each open set of the topology on T under the reverse mapping of f is an open set of the topology on S. Definition 4

Our intuitive ideas about the continuity of real functions are safe, as we have seen, but we now have a definition which is much more general and entirely independent of the concept of distance. Note, further, that the situation is more under our control now: given S, T and $f: S \longrightarrow T$ we can make f continuous or discontinuous by an appropriate choice of topologies.

We consider finally one further example of our ability to generalize well-known ideas.

There is one important idea which seems to depend inextricably on the concept of "nearness": the idea of a limit. (See *Unit 7, Sequences and Limits I.*) Main Text

A sequence of real numbers

$$x_1, x_2, x_3, \ldots$$

(*continued on page 34*)

Solution 1 — **Solution 1**

Let S be the set $\{a, b, c\}$.

The possible topologies are:

$\{\varnothing, S\}$
$\{\varnothing, \{x\}, S\}$, where $x = a, b$ or c,
$\{\varnothing, \{x, y\}, S\}$, where x, y are any two distinct elements of $\{a, b, c\}$,
$\{\varnothing, \{x\}, \{x, y\}, \{y\}, S\}$, $\{\varnothing, \{x\}, \{x, y\}, S\}$, where x, y are any two distinct elements of $\{a, b, c\}$,
$\{\varnothing, \{x, y\}, \{y\}, \{y, z\}, S\}$, $\{\varnothing, \{x\}, \{y, z\}, S\}$, $\{\varnothing, \{x\}, \{y\}, \{x, y\}, \{x, z\}, S\}$, where x, y, z are distinct elements of $\{a, b, c\}$,
the set of *all* subsets of S. ■

(continued from page 33)

is said to have limit a if, given any real number $\varepsilon > 0$, there always exists some positive integer M such that

$$|x_n - a| < \varepsilon \qquad \text{for all } n > M.$$

It is easy to generalize this definition to cope with sequences of points in R^n; all we need do is to replace the modulus by the metric (the distance function). Thus, a sequence of points

$$x_1, x_2, x_3,$$

in R^n will be said to have limit a if, given any real number $\varepsilon > 0$, there always exists some positive integer M such that

$$d(x_n, a) < \varepsilon \qquad \text{for all } n > M.$$

Note that we have already used this generalization in *Unit 28, Linear Algebra IV*, section 28.2.2. We defined the convergence of a sequence of vectors in R^2 by using the concept of a *norm*, which is related to the concept of *distance*.

Let us see if we can re-define this notion in terms which do not include distance.

First we make use of the idea of a neighbourhood. The sequence $x_1, x_2, x_3, \ldots$ of points in R^n has limit a if, for any neighbourhood $N(a, \varepsilon, R^n)$, all the points after some x_M in the sequence belong to the neighbourhood.

Now replace "any neighbourhood $N(a, \varepsilon, R^n)$" by "any open set containing a", and we have the definition we seek. Thus, more generally:

A sequence $x_1, x_2, x_3, \ldots$ of points of a topological space S may be said to have limit a if *every* open set of S containing a also contains all the points of the sequence beyond x_M, where M is some positive integer. — Definition 5

Once again we have generalized a concept in such a way that we can apply it more widely, and yet have not lost the particular situation from which it arose. Restricted to the real sequences we met in *Unit 7*, with open sets interpreted in terms of the usual topology on R, our new definition coincides with the old.

35.4 CONCLUSION

35.4

Conclusion

In this unit we have generalized the concepts of neighbourhood, continuity and limit. We have not done much with the generalizations, mainly because we haven't got the time. But why then bother with the generalizations at all? There are many answers. Without some idea of topology, which is one of the major topics in mathematics today, the Foundation Course would have been incomplete. This unit should have given you some slight impression of the subject.

Another answer is that it exemplifies an important aspect of mathematical thinking and research: the process of generalization. In these elementary aspects, generalization is familiar to all. School algebra, regarded as symbolic arithmetic, is one of our early encounters with generalization. Here we have described a sophisticated example.

Yet another answer is a little more tenuous, but probably more important. As we have said, topology has become a vast and still growing area of mathematics; over the last thirty or forty years it has developed into a well-disciplined subject. Nevertheless it is only recently that it has begun to find application outside mathematics. In the past, the main application of mathematics has been to the exact sciences. In general, when making a mathematical model in the exact sciences, concepts like number and distance are important, because the mathematical analysis required is usually quantitative. But, more recently, mathematics has begun to be applied to other fields such as economics, psychology, etc. A start has been made in trying to model situations of quite a different nature. We can conceive of mathematics dealing with ideas like "conceptual neighbourhoods", that is, constructing a model in which the open sets are made up of related concepts, in the sense in which Polya asks "Can you think of a related problem?"

This may seem very tenuous, and so it is. It may never succeed to any great extent. But we feel that it is entirely appropriate towards the end of a Foundation Course to give some impression of an area which is still developing. Too often one gets the impression when learning mathematics that it is all cut and dried, and that we are transmitting a (dead) body of known facts, instead of studying a living and growing subject.

Acknowledgement

Grateful acknowledgement is made to the following source for illustrations used in this correspondence text:

The Mansell Collection for Leonhard Euler, Gottfried Wilhelm Leibniz, Henri Poincaré.

M100 - MATHEMATICS FOUNDATION COURSE UNITS

1	Functions
2	Errors and Accuracy
3	Operations and Morphisms
4	Finite Differences
5	NO TEXT
6	Inequalities
7	Sequences and Limits I
8	Computing I
9	Integration I
10	NO TEXT
11	Logic I — Boolean Algebra
12	Differentiation I
13	Integration II
14	Sequences and Limits II
15	Differentiation II
16	Probability and Statistics I
17	Logic II — Proof
18	Probability and Statistics II
19	Relations
20	Computing II
21	Probability and Statistics III
22	Linear Algebra I
23	Linear Algebra II
24	Differential Equations I
25	NO TEXT
26	Linear Algebra III
27	Complex Numbers I
28	Linear Algebra IV
29	Complex Numbers II
30	Groups I
31	Differential Equations II
32	NO TEXT
33	Groups II
34	Number Systems
35	Topology
36	Mathematical Structures

SBN 335 01034 2